

Figure 1A

1	CACCAGCAGTAGTAGCAGAAGCGAAGAGCGCAAACGCAACGCTCTCCCGCGCGTTGGC	60
61	CGATTCAATTAATGCAGCTGGCAGCAGAGTTTCCCGACTGGAAAGCGGGCAGTAGCGCA	120
121	ACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTC	180
181	CGGCTCGTATGTTGTGTGGAATTGTGAGCGSATAACAATTTACACAGGAAACAGCTATG	240
241	ACCNTGATTACGCCAAGCTCGAAATTAAACCTCACTAAAGGGAACAAAAGCTGGAGCTCC	300
301	ACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGGAATTCGGCACGA	360
361	GAGGCGCGCGAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCGAGCCAT	420
1		M 1
421	GGCTTCCCTGGGGCAGATCCTCTTCCTGGAGCATAATTAGCATCATCATTATCTGGCTGG	480
2	A S L G Q I L F W S I I S I I I I L A G	21
481	AGCAATTGCACCTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCAGTCACTAC	540
22	A I A L I I I G F G I S G R H S I T V T T	41
541	TGTCGCCCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTTGAACCTGA	600
42	V A S A G N I G E D G I L S C T F E P D	61
601	CATCAAACTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTCAGGCTTGGTCCA	660
62	I K L S D I V I Q W L K E G V L G L V H	81
661	TGAGTTCAAAGAAGGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTTCAGAGGCCGAC	720
82	E F K E G K D E L S E Q D E M F R G R T	101
721	AGCAGTGTGTGCTGATCAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCA	780
102	A V F A D Q V I V G N A S L R L K N V Q	121
781	ACTCAGATGCTGGCACCTACAAATGTTATATCACTTCTAAAGGCAAGGGGAATGC	840
122	L T D A G T Y K C Y I I T S K G K G N A	141
841	TAACCTTGAGTATAAAACTGGAGCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGC	900
142	N L E Y K T G A F S M P E V N V D Y N A	161
901	CAGCTCAGAGACCTTGCGGTGTGAGGCTCCCGATGTTCCCCAGCCACAGTGGTCTG	960
162	S S E E T L R C E A P R W F P Q P T V V W	181
961	GGCATCCCAAGTTGACCAAGGAGCCAACTCTCGGAAGTCTCCAATACCAAGCTTTGAGCT	1020
182	A S Q V D Q G A N F S E V S N T S F E L	201

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1B

1021	GAAC	TCTG	AAGT	GTG	ACCAT	G	AAG	GT	TG	TCT	TG	CT	T	AC	AA	TG	TT	AC	GA	T	CA	AA		1080	
202	N	S	E	N	V	T	M	K	V	V	S	V	L	Y	N	V	T	I	N	N				221	
1081	CAC	ATA	CT	CC	T	GT	AT	G	AT	T	G	A	A	A	A	T	G	C	C	A	A	G	C	A	1140
222	T	Y	S	C	M	I	E	N	D	I	A	K	A	T	G	D	I	K	V	T					241
1141	AGA	AT	CG	G	A	T	C	A	A	A	G	C	G	A	G	T	C	A	C	T	T	A	C	G	1200
242	E	S	E	I	K	R	R	S	H	L	Q	L	L	N	S	K	A	S	L	C					261
1201	TG	CT	CT	CT	TT	CT	TT	GC	CA	T	C	A	G	C	T	T	G	C	T	C	T	C	A	G	1260
262	V	S	S	F	F	A	I	S	W	A	L	L	P	L	S	P	Y	L	M	L					281
1261	AAA	TA	AT	GT	GC	CT	TG	CC	CA	AAAA	A	G	C	A	T	G	C	AA	A	G	T	C	A	1320	
282	K	*																						283	
1321	AGA	AT	T	AT	T	T	C	ACC	ACC	AG	A	T	A	T	G	AC	T	A	G	CT	AT	T	T	1380	
1381	AT	C	T	A	G	A	G	T	C	T	G	G	A	G	T	G	A	G	C	A	A	A	A	1440	
1441	G	C	T	C	A	A	T	A	G	A	T	A	A	T	C	T	A	T	C	T	C	A	A	1500	
1501	T	C	A	T	G	T	G	A	C	T	A	G	A	C	A	A	G	T	G	T	A	A	G	1560	
1561	A	T	C	C	C	C	A	G	A	T	C	T	C	A	G	G	S	A	C	T	C	C	C	1620	
1621	T	G	C	A	T	G	T	C	T	T	G	T	C	T	G	A	A	T	T	T	T	A	G	1680	
1681	C	C	T	G	G	A	A	G	T	C	T	A	T	C	C	C	A	A	T	A	T	C	C	1740	
1741	G	T	A	C	C	T	A	A	G	C	T	G	C	T	A	T	C	G	A	C	T	C	G	1800	
1801	T	A	A	T	G	G	G	T	C	A	A	T	G	A	T	C	A	C	T	T	T	T	A	1860	
1861	A	A	C	T	G	A	C	A	A	T	G	C	A	A	A	A	T	G	A	T	C	A	T	1920	
1921	G	C	G	A	C	C	G	A	T	T	T	A	A	A	A	A	A	C	T	G	A	C	1980		
1981	A	T	T	T	C	A	G	A	T	G	A	T	G	T	C	A	C	C	G	T	G	A	2040		
2041	T	G	G	C	A	T	T	A	T	G	C	A	C	A	A	G	C	T	C	T	C	T	2100		
2101	A	C	C	T	C	A	G	A	T	C	T	A	G	C	A	T	C	T	A	G	C	A	2160		
2161	T	C	T	C	G	G	G	G	A	T	G	T	C	T	G	A	A	C	A	A	T	T	2220		

Figure 1C

2221 CAGTGTCTACTACCAACTAGTGGATAAAGGCCAGGGATGCTGCTCAACCTCTTACCATTGTA 2280

2281 CAGGACGTCTCCCCATTACAACCTACCAATCCGAAGTGTCAACTGTGTCTAGGACTAAGAA 2340

2341 ACCCTGGTTTTGAGTAGAAAAGGGCCTGGAAAGAGGGAGGCCAACAAATCTGTCTGCTTC 2400

2401 CTCACATTAGTTCATTGGCAAATAAGCATTCTGTCTCTTTGGCTGCTGCCCTCAGCACAGAG 2460

2461 AGCCAGAACTCTATCGGCACCAGGATAACATCTCTCACTGAACAGAGTTGACAAGGCCCT 2520

2521 ATGGGAAATGCCCTGATGGGATTATCTTCAGCTTGTGAGCTTCTAAGTTCTTTCCCTTC 2580

2581 ATTCTACCCTGCAAGCCAAGTTCTGTAAGAGAAATGCCGTAGTTCTAGCTCAGGTTTCT 2640

2641 TACTCTGAATTTAGATCTCCAGACCCCTTCTGGCCACAATTCAAATTAAGGCAACAAACA 2700

2701 TATACCTTCCATGAAGCACACACAGACTTTTGAAGCAAGGACAATGACTGCTTGAATTG 2760

2761 AGGCCTTGAGGAATGAAGCTTTGAAGGAAAAGAACTTTGTTTCCAGCCCCCTTCCCAC 2820

2821 ACTCTTCATGTGTTAACCACCTGCCCTTCCTGGACCTTGGAGCCACGGTGACTGTATTACAT 2880

2881 GTTGTATAGAAAACCTGATTTTATAGATTCTGATCGTTCAAGAGAATGATTAATATACAT 2940

2941 TTCTTAAAAAATAAAAAAAAAAACTCGAGGGGGGCCCGGTACCCAATTTCGCCCTATAGT 3000

3001 GAGTCGTATTACAATTCACTGGCCGTGTTTTTACAACGTCTGACTGGGAAAACCCCTGGC 3060

3061 GTTACCACACTTAATCGCCTTGCAGCAGATCCCCCTTTGCAGCTGGCGTAATAGCGAA 3120

3121 GAGGCCCGCACCGATCGCCCTTCCCAACAKTTGCGCAGCTGAATGGCGAATGGCAAATT 3180

3181 GTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTTTT 3240

3241 AACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAAATAGACCGAGATAGGG 3300

3301 TTGAGTGTGTTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGTGTTCACCGCGGTGA 3357

Figure 2

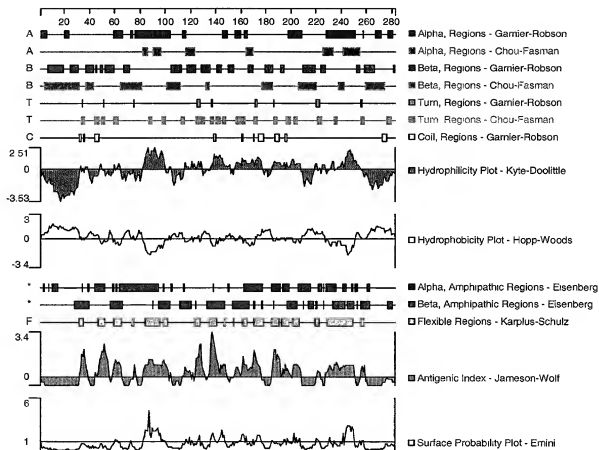


Figure 3A

1	CCACGCGTCCGGAATGAACAACTTTTCTCTCTTGAATATATCTTAACGCCAAATTTTGA	60
61	GTGCTTTTTTGTATTACCATCCTCATATGTCCAGCTGGAAAGAATCCTGGGTGGAGCTA	120
121	CTGCATGTTGATTGTTTTGTTTTCTCTTTGGCTGTTTCATTTTGGTGGCTACTATAAGGA	180
181	AATCTAACACAAACAGCAACTGTTTTTGTGTTTACTTTTGCATCTTTACTTGTGGAGC	240
241	TGTGGCAAGTCTCATATCAAATACAGAACTGATCTTCTCCTGCATAAGTTTGAGCCCTG	300
1	M I F L L L M L S L	10
301	GAATGACAGCTTCACCAGATAGCAGCTTTATTACAGTGACATCCCTAGGAACTGTAC	360
11	E L Q L H Q I A A L F T V T V P K E L Y	30
361	ATAATAGAGCATGGCAGCAATGTGACCCCTGGAATGCAACTTTGACACTGGAAGTCATGTG	420
31	I I E H G S N V T L E C N F D T G S H V	50
421	AACCTTGGAGCAATAACAGCCAGTTTGCAAAAGGTGGAAATGATACATCCCCACACCTG	480
51	N L G A I T A S L Q K V E N D T S P H R	70
481	GAAAGGCCACTTTGCTGGAGGAGCAGCTGCCCTAGGGAAGGCTCGTTCACATACCT	540
71	E R A T L L E E Q L P L G K A S F H I P	90
541	CAAGTCCAAGTGAGGGACGAAGGACAGTACCAATGCATAATCATCTATGGGGTACGCCTGG	600
91	Q V Q V R D E G Q Y Q C I I I Y G V A W	110
601	GACTACAAGTACCTGACTCTGAAAGTCAAAGCTTCTTACAGGAAAATAAACACTCACATC	660
111	D Y K Y L T L K V K A S Y R K I N T H I	130
661	CTAAAGGTTCCAGAAACAGATGAGGTAGAGCTCACCTGCCAGGCTACAGGTATCTCTCTG	720
131	L K V P E T D E V E L T C Q A T G Y P L	150
721	GCAGAAGTATCTGGCCAAACGTCAGCGTTCTGCCAACACCAGCCACTCCAGGACCCCT	780
151	A E V S W P N V S V P A N T S H S R T P	170
781	GAAGGCCTCTACCAGGTACACAGTGTGTTGCGCCTAAAGCCACCCCTGGCAGAACTTC	840
171	E G L Y Q V T S V L R L K P P P G R N F	190
841	AGCTGTGTGTTCTGGAATACTACGTGAGGGAACCTTACTTTGGCCAGCATTGACCTCAA	900
191	S C V F W N T H V R E L T L A S I D L Q	210
901	AGTCAGATGGAACCCAGGACCCATCCAACCTTGGTGCTTCACATTTTCATCCCCCTCCTGC	960
211	S Q M E E P R T H P T W L L H I F I P S C	230
961	ATCATTGCTTTTCATTTTCATAGCCACAGTGATAGCCCTAAGAAAAACACTCTGTCAAAG	1020
231	I I A F I F I A T V I A L R K Q L C Q K	250

Figure 3B

1021	CTGTATTCTTCAAAGACACAACAAAAAGACCTGTACCCACAACAAAGAGGGAAGTGAAC	1080
251	L Y S S K D T T K R P V T T T K R E V N	270
1081	AGTGTCTGTGAATCTGAACCTGTGGTCTTGGGAGCCAGGGTGACCTGATATGACATCTAAA	1140
271	S A V N L N L W S W E P G *	284
1141	GAAGCTTCTGGACTCTGAACAAGAATTCGGTGGCCTGCAGAGCTTGCCATTGCACTTTT	1200
1201	CAAATGCCTTTGGATGACCCAGCACTTTAATCTGAAACCTGCAACAAGACTAGCCAAACAC	1260
1261	CTGGCCATGAAACTTGCCCCCTCAGCTGATCTGGACTCACCTCTGGAGCCTATGGCTTTAA	1320
1321	GCAAGCACTACTGCACCTTTACAGAATTACCCCACTGGATCCTGGAGCCACAGAATTCCTT	1380
1381	CAGGATCCTTCTTGCTGCCAGACTGAAAGCAAAGGAATTTATTTCCCTCAAGTTTCTTA	1440
1441	AGTGATTTCCAAAGCAGAGGTGTGTGGAATTTCCAGTAACAGAAACAGATGGGTGGC	1500
1501	AATAGAGTTATTTTTTATCTATAGCTTCCCTCGGGTACTAGAAGAGGCTATTGAGACTAT	1560
1561	GAGCTCACAGACAGGGCTTCGCACAAACTCAAATCATAATTGACATGTTTTATGGATTAC	1620
1621	TGGAATCTTGATAGCATAATGAAGTTGTTCTAATTAACAGAGAGCATTTAAATATACACT	1680
1681	AAGTGCACAAATTGTGSAGTAAAGTCATCAAGCTCTGTTTTTGTAGGTCTAAGTCACAAAG	1740
1741	CATTGTTTTAACCTGTAATGGCACCATGTTTAATGGTGGTTTTTTTTTGAACATACATC	1800
1801	TTTCCTTTAAAAATTATTTGGTTCTTTTTATTGTTTTTACCTTAGAAAATCAATATATATA	1860
1861	CAGTCAAAAATATTTGATATGCTCATACGTTGTATCTGCAGCAATTCAGATAAGTAGCT	1920
1921	AAAATGGCCAAAGCCCCAAACTAAGCCTCCTTTTCTGGCCCTCAATATGACTTTAAATTT	1980
1981	GACTTTTCAGTGCCTCAGTTTGCACATCTGTAATACAGCAATGCTAAGTAGTCAAGGCCT	2040
2041	TTGATAATTGGCACTATGGAAATCCTGCAAGATCCCACTACATATGTGTGGAGCAGAAGG	2100
2101	GTAACCTCGGCTACAGTAACAGCTTAATTTGTTAAATTGTTCTTTATACATGGAGCCATG	2160
2161	AAGCTCAGAGCATTAGCTGACCCCTGAACTATTCAAATGGGCACATTAGCTAGTATAACA	2220
2221	GACTTACATAGGTGGGCTAAAGCAAGCTCCTTAACTGAGCAAAATTTGGGGCTTATGAG	2280

**Figure 3C**

2281 AATGAAAGGGTGTGAAATTGACTAACAGACAAATCATACATCTCAGTTTCTCAATTCTCA 2340  
2341 TGTAAATCAGAGAATGCCTTTAAAGAATAAACTCAATTGTTATTCTTCAAAAAAAAAA 2400  
2401 AAAAAA 2406

1000  
900  
800  
700  
600  
500  
400  
300  
200  
100  
0  
1000  
900  
800  
700  
600  
500  
400  
300  
200  
100  
0

Figure 4

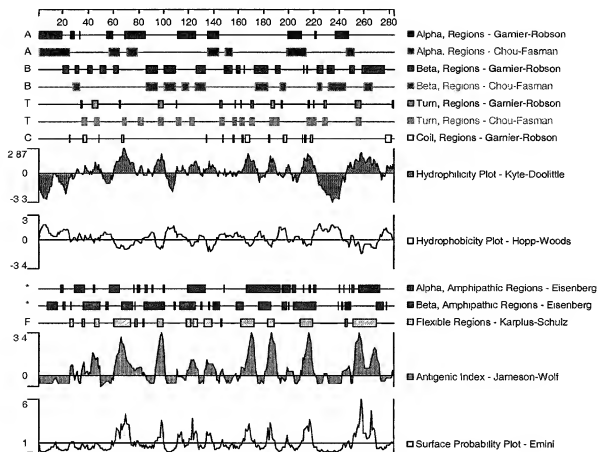




Figure 5A

1	GGCAGAGCTGTCATCCGTTTCCATGCCGTGAGGTCCATTCCACAGAACACATCCATGGCT	60
1		2
	M A	
61	CTCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGT	120
3	L M L S L V L S L L K L G S G Q W Q V F	22
121	GGCCACAGCAAGCCTGTCCAGGCCCTGTGTTGGGGAGGACGCAGCATTCTCTGTCTCTCG	180
23	G P D K F V Q A L V G E D A A F S C F L	42
181	TCTCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGTCTCTTAGC	240
43	S P K T N A E A M E V R F F R G Q F S S	62
241	GTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTCAGCAGATGCCACAGTATCAA	300
63	V V H L Y R D G K D Q P F M Q M P Q Y Q	82
301	GGCAGGACAAAACGTGGAAGGATTCCTATTCGGAGGGGCCATCTCTCTGAGGCTGGAA	360
83	G R T K L V K D S I A E G R I S L R L E	102
361	AACATTACTGTGTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGCTCTTACTAC	420
103	N I T V L D A G L Y G C R I S S Q S Y Y	122
421	CAGAAGGCCATCTGGGAGCTACAGGTGTGAGCACTGGGGTCAGTTCTCTCATTTCCATC	480
123	Q K A I W E L Q V S A L G S V P L I S I	142
481	GCGGATATGTTGATAGAGACATCCAGTACTCTGTGAGTCTTCGGGCTGGTTCCTCCCGG	540
143	A G Y V D R D I Q L L C Q S S G W F P R	162
541	CCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTTGTCACAGACTCCAGGACAAAC	600
163	P T A K W K G P Q G Q D L S T D S R T N	182
601	AGAGACATGCATGCCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAACGCCGG	660
183	R D M H G L F D V E I S L T V Q E N A G	202
661	AGCATATCTGTTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGTACAG	720
203	S I S C S M R H A H L S R E V E S R V Q	222
721	ATAGGAGACTGGAGAAGAAAGCAGGACAGGCAGGTAAAGAAAAATATTCTCTTCACAC	780
223	I G D W R R K H G Q A G K R K Y S S S H	242
781	ATTATGACTCTCTTCCAGTCTCTCGTTTATGGATTTTATATCCTGAGGCCCGTGGGT	840
243	I Y D S F P S L S F M D F Y I L R P V G	262
841	CCCTGCAGAGCCAAGCTTGTGATGGGAACCTGAAATTGACAGATTCTGGGGGAGGTGCAT	900
263	P C R A K L V M G T L K L Q I L G E V H	282
901	TTTGTAGAGAAGCCCCATAGCCTTCTCAGATCTCTGGAGGGTCCACAACACTCAAAAG	960
283	F V E K P H S L L Q I S G G S T T L K K	302

Figure 5B

961	GGTCCCAATCCTTGGTCTTTCCCTTCTCCCTGCGCCCTGTTTCCACGTGAGCACGGAAC	1020
303	G P N P W S F P S P C A L F P T *	319
1021	TGCCCTGCTCTCTCTGCTTGCTTTCAGAAATTGAGAGACGCCCGAAACACGCAGGTACCAA	1080
1081	CGCCTGAGAGGTTAACAGTGGGCATGGAGTAGGAAGATGACCAGTGACAGATATGGAGCC	1140
1141	CATCCAGCTTGTAGACAGCAARTCTGTGATGCCCGAATCCACCCAGGGTGCAGCTGCCT	1200
1201	CTAAATACACTTCTTGGCCCAGGACTTGGAGGGAAAAGCGTAGGGACTGGGTCACTAGG	1260
1261	AGGGGTACAGGCAAGACGCCAGGGAACTAGGGCATTAGTAGCTGGCTTCTAGGGTCT	1320
1321	GTGCAAGGGGAACGAAGTGAAGTTAGCAGGAACTGGTGGGTGGAAGGAAGCTGAATCCT	1380
1381	GGAGTCACTCAAGGTCTCACAAAGTCAAATAGAGGGCTTACGTGGGAGGGCAGTGGTAGG	1440
1441	GCTGGGTGAACATCTCATGGTTGAGCATCTCCAAGCATCAGTGAGGCACGGGGGCTGCC	1500
1501	TGGAGAAGGTACATGGCTGGTGGGATAGTGGGACTGGCCGGATCCTACCCGGAGCCAGTC	1560
1561	TGCAGTGGGAGGGTCGACCTCTTGCTCCAGCCAGATTTCGTCTTCAGTAACATCATGCTT	1620
1621	CCTCTCTCCCCACCGCACCCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGA	1680
1681	AGCTCTGCGTTTCTGATCTGAAAACGTAAACCATAGAAAAGCTCCTCAGGAGGTGCCTC	1740
1741	ACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTTCAGGGTTTCCAAGCAGGGA	1800
1801	AACATTACTGGGAGGTGGACGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTGGG	1860
1861	ATGACGTAGACAGGGGAAGAACAATGTGACTTTGTCTCCCAACAATGGGTATTTGGGTCC	1920
1921	TCAGACTGACAACAGAACATTGTATTTACATTCGAATCCCATTTTATCAGCCTCCCC	1980
1981	CCAGCACCCCTCCTACACGAGTAGGGGTCTTCTGGACTATGAGGGTGGGACCATCTCCT	2040
2041	TCCTTCAATCAAAATGACCAGTCCCTTATTTATACCTGCTGACATGTGAGTTTGAAGGT	2100
2101	TGTTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAGGGGACTCCCATATTCA	2160
2161	TATGTCCAGTGTCTTGGGATGAGACAGAGAAGACCTGCTTAAAGGGCCCCACACCACA	2220

Figure 5C

2221 GACCCAGACACAGCCAAGGGAGAGTGC<sup>1</sup>TCCGACAGGTGGCCCCAGCT<sup>2</sup>TCCTCTCCGGAG<sup>3</sup> 2280

2281 CCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTC<sup>4</sup>TGCCCT<sup>5</sup> 2340

2341 GAGCCCTGCAGCAGCGGCAGTCA<sup>6</sup>CAGCTTCCAGATGAGGGGGATTGGCTGACCC<sup>7</sup>TGTG<sup>8</sup> 2400

2401 GGAGTCAGAAGCCATGGCTGCCCTGAA<sup>9</sup>GTGGGACGGAATAGACTCACATTAGGTTT<sup>10</sup>TAGT<sup>11</sup> 2460

2461 TTGTGAAAACTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCA<sup>12</sup> 2520

2521 TTTGCTAGTCACGGACAGTGATTCC<sup>13</sup>TGCC<sup>14</sup>TCACAGGTGAAGATTAAAGAGACAACGAATG<sup>15</sup> 2580

2581 TGAATCATGCTTGCAGGTTTGAGGGCCACAGTGT<sup>16</sup>TTGCTAATGGATGTGTTTTTATGATT<sup>17</sup> 2640

2641 ATACATTTTCCCCACCATAAACTCTGTTTGCCCTTAATTCCCACATTAA<sup>18</sup>TTAACTTTTC<sup>19</sup> 2700

2701 CTCTTATACCCAAATCCACCATGGAATAGTTA<sup>20</sup>ATTGGAACACCTGCCTTTGTGAGGCTC<sup>21</sup> 2760

2761 CAAAGAATAAAGAGGAGGTAGGATTTT<sup>22</sup>CTACTGATTCTATAAGCCAGCATTTACCTGATA<sup>23</sup> 2820

2821 CCAAAACAGGCAAGAAAAACAGAAGAAGAGGAAGGAAAACTACAGGTCCATATCCCTCA<sup>24</sup> 2880

2881 TTAACACAGACACAAAAATCTAAATAAAATTTTAA<sup>25</sup>CAAAATTAACTAAACAATATATTT<sup>26</sup> 2940

2941 AAAGATGATATATAACTACTCAGTGTG<sup>27</sup>TTGTCCCACAAATGCAGAGTTGGTTTAATAT<sup>28</sup> 3000

3001 TTAAATATCAACCAGTGTAA<sup>29</sup>TTCAGCACATTAATAAAGTAAAAA<sup>30</sup>AAAAAAAAAAAAAAAAAAAAA 3059

111  
122  
133  
144  
155  
166  
177  
188  
199  
200  
211  
222  
233  
244  
255  
266  
277  
288  
299  
300

Figure 6

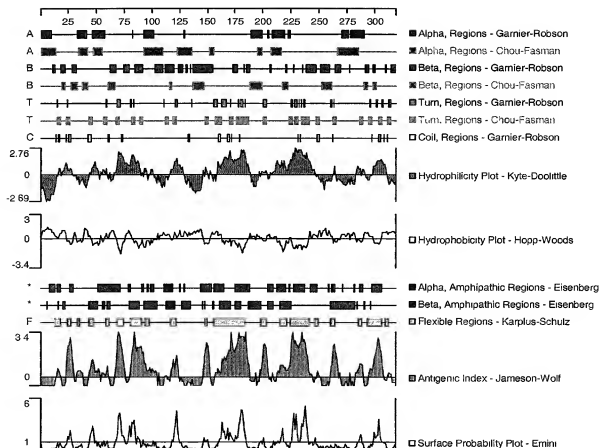


Figure 7A

1 NNACAGAGCCTGTGCCCTGGAAAGGTTGGAGACTTGGGGGACGACTGGAGAATTGCCAT 60

61 TTGAGGACCAAAGGAGAAAAGAACTACACGCTAATTCTAGAAGGCCTCCTGTCCCTGCC 120

121 TGCTCTGGGTGCTCATGGAACAGCTGCTGCCCTGCACTTCTCCCGGCCAGCCTCCCTCC 180  
1 M E P A A A L H F S R P A S L L 16

181 TCCTCCTCTCAGCCTGTGTGCACTGGTCTCAGCCCAAGTTTACTGTCTGGGGCCAGCTA 240  
17 L L L S L C A L V S A Q F T V V G P A N 36

241 ATCCCATCTGGCCATGGTGGGAGAAAACACTACATTACGCTGCCATCTGTACCCGAGA 300  
37 P I L A M V G E N T T L R C H L S P E K 56

301 AAAATGCTGAGGACATGGAGGTGCGGTGGTTCGGGTCTCAGTTCTCCCGCAGTGTGTTG 360  
57 N A E D M E V R W F R S Q F S P A V F V 76

361 TGTATAAGGGTGGGAGAGAGAAACAGAGGAGCAGATGGAGGAGTACCGGGGAAGAATCA 420  
77 Y K G G R E R T E E Q M E E Y R G R I T 96

421 CCTTGTGAGCAAAAGACATCAACAGGGGCAGCGTGGCCCTGGTCATACATAACGTCACAG 480  
97 F V S K D I N R G S V A L V I H N V T A 116

481 CCCAGGAGAATGGGATCTACCGCTGTTACTTCCAAGAAGGCAGGTCTACGATGAGGCCA 540  
117 Q E N G I Y R C Y F Q E G R S Y D E A I 136

541 TCCTACGCCCTCGTGGTGGCAGGCCTTGGGTCTAAGCCCTCATTTGAAATCAAGGCCAAG 600  
137 L R L V V A G L G S K P L I E I K A Q E 156

601 AGGATGGGAGCATCTGGCTGGAGTGCAATCTGGAGGGTGGTACCCAGAGCCCTCACAG 660  
157 D G S I W L E C I S G G W Y P E P L T V 176

661 TGTGGAGGGACCCCTACGGTGAGGTGTGTCGCCGCCCTGAAGGAGGTTTCATCGCTGATG 720  
177 W R D P Y G E V V P A L K E V S I A D A 196

721 CTGACGCCCTCTTCATGGTCAACACAGCTGTGATCATCAGAGACAAGTATGTGAGGAATG 780  
197 D G L F M V T T A V I I R D K Y V R N I 216

781 TGTCTGCTCTGTCAACAACACCCCTGCTCGGCCAGGAGAAGAACTGTCAATTTTATTC 840  
217 S C S V N N T L L G Q E K E T V I F I P 236

841 CAGAATCCTTTATGCCAGCGCATCTCCCTGGATGGTGGCCCTAGCTGTCACTCCTGACCG 900  
247 E S F M P S A S P W M V A L A V I L T A 256

901 CATCTCCCTGGATGGTGTCCATGACTGTATCCTGGCTGTTTTTCATCATCTTATGGCTG 960  
257 S P W M V S M T V I L A V F I I F M A V 276

Figure 7B

961	TCAGCATCTGTTGCATCAAGAAACTTCAAGGGGAAAAAAGATTCTGTACAGGGGAAAAAGA	1020
277	S I C C I K K L Q R E K K I L S G E K K	296
1021	AAGTTGAACAAGAGGAAAAAAGAAATGCACAGCAACTTCAAGAAGAAATGCGATGGAGAA	1080
297	V E Q E E K E I A Q Q L Q E E L R W R R	316
1081	GAACATTCTTACATGCTGCTGATGTGGTCTGGATCCAGACACCGCTCATCCGAGCTCT	1140
317	T F L H A A D V V L D P D T A H P E L F	336
1141	TCCTGTCAGAGGACCGGAGAAAGTGTGAGGCGGGGCCCCACAGGCAGAGAGTGCCTGACA	1200
337	L S E D R R S V R R G P Y R Q R V P D N	356
1201	ACCCAGAGAGATTGACAGTCAGCCTTGTGTCTCTGGATGGGAGAGCTTGCCTCAGGGA	1260
357	P E R F D S Q P C V L G W E S F A S G K	376
1261	AACATTACAGGGGAAACTTCAAGAGTGGGGACCCACAGAGCCTATAGAAATCAATTCCT	1320
377	H Y R G N F T E W G P T R A Y R I N S L	396
1321	TGGACTCACAGCCATGCAGAAAGCCCTGGCCATCTCAGCAGCCACCGACAAACCCCTTA	1380
397	D S Q P C R K P W P S Q Q P P H N P P N	416
1381	ATGAAAGACACGCCCTCCCTCCCTCTGGTCACGTAAGAAACATCTTCCAGCTGCCTTTT	1440
417	E R H A L L P S G H V R E H L P A A F F	436
1441	TCACACCCACTCCAGCCCCTGCCCCAGTTTCTCCTCCTCACTAGTCTGTGGCTTTAGT	1500
437	T P T P A L C P S F L L L T S L W L *	455
1501	AGTTCCTTTGCTTGTAAATTAATGGGATGGGATCCAGGCATAGGGAACATAGTTGTTTCATAG	1560
1561	CTCCCACTCAAAAAGAAAGTGAGAGAAGCTGTTGGGCAGCGAACCTACTGTTTAAATCA	1620
1621	GGATAACCACTTAAGCCCAATATGCCAGTTGGCACCAGATGCTGTGGAGATTGGAATGAG	1680
1681	GCCAACAGGGTTTACCAGGATGAGAGAGGAGAGAGGAATCCACAGGACCACAGAAAGGA	1740
1741	GAGGGAACCATATGAGATCAGAGATAGAGGAAGTGTGTGAGAGGAAAGGGGAGGTCT	1800
1801	GCTGATTCCCTCAGAAATGGCTTCTGGACCCTGGAGATGTTTGGAAACCAATACCGGGCCCT	1860
1861	GTCCCTCCCTGAGAGGATTCTCCCTTTGAAGGAGTCCCTTTGCCGGGTGGGCGTCTTCCT	1920
1921	GGACTATGAAGCTGGAGATGTCTCTTCTACAACATGAGGACAGATCACACATCTACAC	1980
1981	ATGTCCCCGTTACGCCTTTAATGTGCCTGTGAGGCCATCTTCAGGTTCAGGTCGTGATGA	2040

bioRxiv preprint doi: <https://doi.org/10.1101/100000>; this version posted April 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 7C

2041 CAGCCCCATCTTCATCTGCCTGCACTCACAGGAGCCAGTGGGGTCATGGTGCCTGAAGA 2100  
2101 GGGCCTGAAACTTCA CAGAGTGGGACCCACCAAGGTTGTAAGGATGGCTAAGTCCCACC 2160  
2161 ATAAGAGCTAAAGGGTCCTGGGAGATGATGGCTCATTTCCACCCCAACCCAGGATTCCA 2220  
2221 CAGCACACACCCACAGGCCTGGACCTGGGATGAAGATGAATGAAGAACATGGACTCATGT 2280  
2281 GGATGTGGTTTGGCTCAGATGTCCCTGCAATAACAAGGGTCAGTACTTAGTCCCTGAG 2340  
2341 TGTGGTTGAGGTTTGAGGTCCTGGTCGAGCAGGGCAGTACTGGAC CAGGTCTACGTCAGC 2400  
2401 ATT CAGGTTCAATGGGGACAC CAGTGGCTTCAAACCTCCTGATCTAATTA TGT TTTTAGA 2460  
2461 CACTTAGAAGTTATTGAGGACTTTAAAGAACTTTTGTTTATTTGGGTTAATATTATGAC 2520  
2521 ATTTGACCATTGAAACAAAAATTTAAATGTTATCTTTTAATTTATGTTAAATAGCATT 2580  
2581 AATAAATCAGTTATAGGTTAATGTAGATAGGATGTTTTGTGAAAAAGCAATCTATTGTGT 2640  
2641 CCAAATAAAAAACAAAAAGTGTA AAAAAAAAAAAAAAAAAAAAA 2682

Figure 8

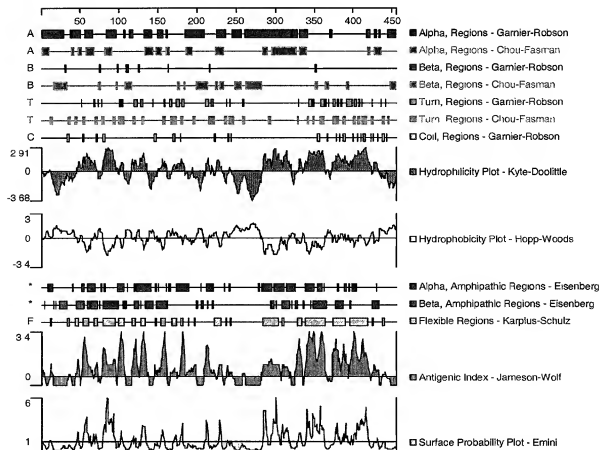




Figure 9A

1	CGATTCCGCTCCAAACTCCGGCGCTGCAGCCGATCGGACTCTGGGCCCGCGGTGGGCACCG	60
61	CGCGCAGCTAGGGAGCCGAGAAACCGCGCGAGCCCCGAGGACGCCAGAGCGCGAGGGTC	120
121	GCTGCGCCTCGCAGAGCCGAGCGAGTCGAGCCGGGCGCCCCGGGCTGCCGTGGAGACGCC	180
181	GTGACTTTGAAGTGTAACTTCAAGACAGATGGGCGCATGCGGGAGATCGTGTGGTACCGG	240
1	M R E I V W Y R	8
241	GTGACGGATGGTGGCACCATCAAGCAAAAGATCTTCACCTTCGACGCCATGTCTCCACC	300
9	V T D G G T I K Q K I F T F D A M F S T	28
301	AACTACTCACACATGGAGAACTACCGCAAGCGAGAGGACCTGGTGTACCAGTCCACTGTG	360
29	N Y S H M E N Y R K R E D L V Y Q S T V	48
361	AGGTCGCCCGAGGTCCGGATCTCAGACAAATGGTCCCTATGAGTGCCATGTGGGCATCTAC	420
49	R L P E V R I S D N G P Y E C H V G I Y	68
421	GACCGCGCCACCAGGAGAAAGTGGTCTCGCATCAGGCAACATCTTCTCAACGTCATG	480
69	D R A T R E K V V L A S G N I F L N V M	88
481	GCTCTCCCACTCCATTTGAAGTGGTGGCTGCTGACACACAGCCCCCTTCAGCCGCTAC	540
89	A P P T S I E V V A A D T P A P F S R Y	108
541	CAAGCCCAAGAACTTCACGCTGGTCTGCATCGTGTCTGGAGGAAAACAGCACCCATGGTT	600
109	Q A Q N F T L V C I V S G G K P A P M V	128
601	TATTTCAAACGAGATGGGGAAACCAATCGACGCAGTGGCCCTATCAGAGCCACCAGCTGCG	660
129	Y F K R D G E P I D A V P L S E P P A A	148
661	AGTCCGGCCCCCTACAGGACAGCAGGCCCTTCCGCAGCCTTCTGCACCGTGACCTGGAT	720
149	S S G P L Q D S R P P R S L L H R D L D	168
721	GACACCAAGATGCAGAAATCACTGTCCCTCCTGGACCCGAGAACCGGGGTGGGCGACCC	780
169	D T K M Q K S L S L L D A E N R G G R P	188
781	TACACGGAGCGCCCCCTCCCGTGGCCTGACCCAGATCCCAACATCTCTCCAGCCCAACC	840
189	Y T E R P S R G L T P D P N I L L Q P T	208
841	ACAGAGAACATACCAGAGACGGTGTGAGCCGTGAGTTTCCCCGCTGGGTCCACAGCGCC	900
209	T E N I P E T V V S R E F P R W V H S A	228
901	GAGCCCACTACTTCTTGGCCACAGCGCGACCCCGAGCAGTGACGGCATGTGTGAAGTA	960
229	E P T Y F L R H S R T P S S D G T V E V	248
961	CGTGCCTGTCTACCTGGACCCCTCAACCCACAGATCGACAACGAGGCCCTCTTCAGCTGC	1020
249	R A L L T W T L N P Q I D N E A L F S C	268

[illegible][illegible]

Figure 10

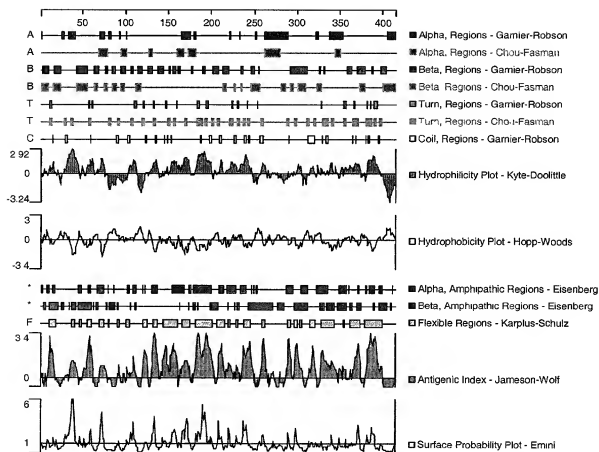


Figure 11

1 CACGAGCCTGTGCCCTGGAAAGGTTGGAGACTTGGGGGACGACTGGAGAAATGCCATTT 60

61 GAGGACCAAAGGAGAAAAAAGAACTACACGCTAATTCTAGAAGCGCTCTGTCCCTGCCTG 120

121 CTCTGGGTGCTCATGGAAACAGCTGTGCTGCCCTGCACTTCTCCCGCCAGCCTCCCTCCTC 180  
1 M E P A A A L H F S R P A S L L 16

181 CTCTCTCTAGCCTGTGTGCACTGGTCTCAGGCCAGGTCACTGTCTGGGGCCCACTGAT 240  
17 L L L S L C A L V S A Q V T V V G P T D 36

241 CCCATCTGGCCATGGTGGGAGAAAACTACGTTACGATGCTGTCTGTCAACCCGAGGAA 300  
37 P I L A M V G E N T T L R C C L S P E E 56

301 AATGCTGAGGACATGGAGGTGCGGTGGTTCAGTCTCAGTTCTCCCTGCACTGTTTGTG 360  
57 N A E D M E V R W F Q S Q F S P A V F V 76

361 TATAAGGTTGGAAGAGAGAGAAACAGAGGAGCAGAGGAGGAGTACCGAGGGAGAACCC 420  
77 Y K G G R E R T E E Q K E E Y R G R T T 96

421 TTTGTGAGCAAAGACAGCAGGGGGCAGCGTGGCCCTGATACACAATGTACAGCCGAG 480  
97 F V S K D S R G S V A L I I H N V T A E 116

481 GATAACGGCATCTACCACTGTTACTTCCAAGAAGGCAGGTCTGCAATGAGGCCATCCTG 540  
117 D N G I Y Q C Y F Q E G R S C N E A I L 136

541 CACCTTGTGGTGGCAGACCAGCACAACTCTCTTTCTCTGGATCCCCATTCGCGAGGGGACA 600  
137 H L V V A D Q H N P L S W I P I P Q G T 156

601 CTCTCCCTATGAAAAGAAGATTCCAGGGGAAAAATCCTTCCCTCTGCACAAGGGCCACCA 660  
157 L S L \* 160

661 TGAGTGAGTTTGCCCTGCTAAGCCGTGGGCTTGACTTCTTGAGAAGCACATGCAGAACTC 720

721 AGTTGAGGCCATGAGCCGGGGGAAAAATGGTGAATCTCGGAAGAGAAGCTCTATGCCTGCC 780

781 TTAGCACTGAGCTGTGCACTTCTGAGAGTGAGAGGAGACACCATCAATAATTGTCTTGGG 840

841 ACAACTGGAATAAACAGTACTGCCCCAGAGAACTACGATATTTGAAATCTTATTCTTGA 900

901 TGAATATTCTCTGACTTCTTTCTCTGAAATGCTGTTTGCAAAGAGAGTGACTTATATGT 960

961 AAGTAGAGCCTTTTATTAAAGCAAGACTTAATACAGAAGCAAAAAAAAAAAAAAAAAA 1019

Figure 12

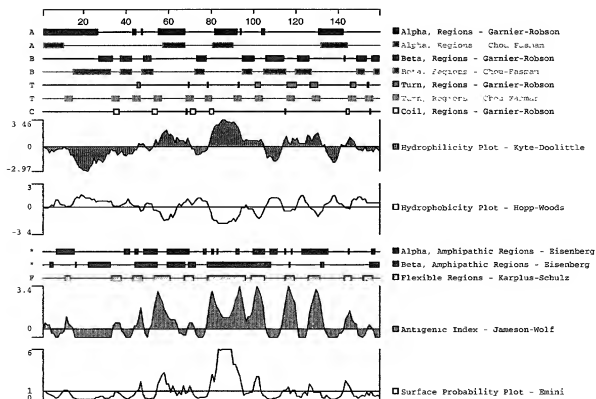


Figure 13A

1	ACATCCATGGCTCTAATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAG	60
1	M A L M L S L V L S L L K L G S G Q	18
61	TGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAGGACGCAGCATTC	120
19	W Q V F G P D K P V Q A L V G E D A A F	38
121	TCCTGTTTCCCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCAGTCTCTCAGGGGC	180
39	S C F L S P K T N A E A M E V R F F R G	58
181	CAGTTCTCTAGCTGGTCCACCTCTACAGGGACGGGAAGGACCAAGCATTATGCAGATG	240
59	Q F S S V V H L Y R D G K D Q P F M Q M	78
241	CCACAGTATCAAGGCAGGACAAAACCTGGTGAAGGATTCTATTCGGAGGGGCGCATCTCT	300
79	P Q Y Q G R T K L V K D S I A E G R I S	98
301	CTGAGGCTGGAAAAACATTACTGTGTTGGATGCTGGCCCTCTATGGGTGCAGGATTAGTTCC	360
99	L R L E N I T V L D A G L Y G C R I S S	118
361	CAGTCTTACTACCAGAAGGCCATCTGGGAGCTACAGGTTCTCAGCACTGGGCTCAGTTCT	420
119	Q S Y Y Q K A I W E L Q V S A L G S V P	138
421	CTCATTTCCATCAGGGATATGTTGATAGAGACATCCAGCTACTCTGTCTAGTCTCTCGGGC	480
139	L I S I T G Y V D R D I Q L L C Q S S G	158
481	TGGTTCCCCCGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTGTCCACAGAC	540
159	W F P R P T A K W K G P Q G Q D L S T D	178
541	TCCAGGACAAACAGAGACATGCATGGCCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAA	600
179	S R T N R D M H G L F D V E I S L T V Q	198
601	GAGAACGCCGGGAGCATATCTGTTCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAA	660
199	E N A G S I S C S M R H A H L S R E V E	218
661	TCCAGGGTACAGATAGGAGATACCTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAA	720
219	S R V Q I G D T F F E P I S W H L A T K	238
721	GTACTGGGAATACTCTGCTGTGGCCTATTTTGGCATTGTTGGACTGAAGATTCTCTTC	780
239	V L G I L C C G L F F G I V G L K I F F	258
781	TCCAAATTCAGTGGAAATCCAGGCGGAACCTGGACTGGAGAAGAAAGCACGGACAGGCA	840
259	S K F Q W K I Q A E L D W R R K H G Q A	278
841	GAATTGAGAGACGCCCGGAAACACGCACTGGAGGTGACTCTGGATCCAGAGACGGCTCAC	900
279	E L R D A R K H A V E V T L D P E T A H	298
901	CCGAAGCTCTGCGTTTCTGATCTGAAAACCTGTAACCCATAGAAAAGCTCCCCAGGAGGTG	960
299	P K L C V S D L K T V T H R R K A P Q E V	318

**Figure 13B**

951	CCTCACTCTGAGAAAGAGATTACAAGGAAGAGTGTGGTGCCTCTCTCAGAGTTTCCAAGCA	1020
319	P H S E K R F T R K S V V A S Q S F Q A	338
1021	GGGAAACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCTGGGAGTGTGC	1080
339	G K H Y W E V D G G H N K R W R V G V C	358
1081	CGGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTTTGCTCCCGATCATGGGTACTGG	1140
359	R D D V D R R K E Y V T L S P D H G Y W	378
1141	GTCTCTCAGACTGAATGGAGAAACATTTGTAATTTCACATTAAATCCCCTGTTTATCAGCGTC	1200
379	V L R L N G E H L Y F T L N P R F I S V	398
1201	TTCCCCAGGACCCACCTACAAAAATAGGGGCTTTCCTGGACTATGAGTGTGGGACCATC	1260
399	F P R T P P T K I G V F L D Y E C G T I	418
1261	TCCTTCTTCAACATAAATGACCAGTCCCTTATTATACCTTGACATGTCCGGTTGAAGGC	1320
419	S F F N I N D Q S L I Y T L T C R F E G	438
1321	TTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAAATGGAATCCCAGAGAC	1380
439	L L R P Y I E Y P S Y N E Q N G T P R D	458
1381	AAGCAACAGTGTGATCCTCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAATGTGA	1440
459	K Q Q *	462
1441	GGATGAATCCATCCACATTCTCTTTTAGGGATATTAAGTCTCTCTCCAGATCCAAA	1500
1501	GTCCCGCAGCAGCCGCCAAGTGGCTTCAGATGAAGGGGAGCTGGCCTGTCCACATGG	1560
1561	GAGTCAGGTGTGATGGCTGCCCTGAGCTGGGAGGGAAGAGGCTGACATTACATTTAGTT	1620
1621	TGCTCTCACTCCATCTGGCTAAGTGATCTTGAATACCACTCTCAGGTGAAGAACCGTC	1680
1681	AGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAAGTAGACAAGGAATGTGAATAATGC	1740
1741	TTAGATCTTATTGATCACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTTCA	1800
1801	TA	1833

Figure 14

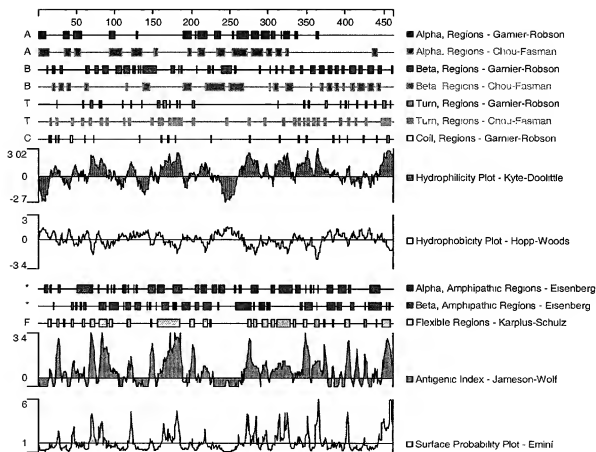




Figure 15

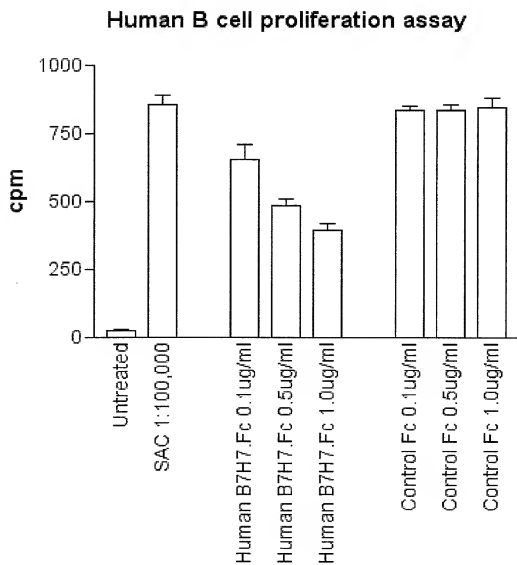


Figure 16

